



Article

A Comparative Analysis of Optimization Algorithms for Gastrointestinal Abnormalities Recognition and Classification Based on Ensemble XcepNet23 and ResNet18 Features

Javeria Naz ¹, Muhammad Imran Sharif ¹, Muhammad Irfan Sharif ², Seifedine Kadry ^{3,4,5,6,*}, Hafiz Tayyab Rauf ⁷ and Adham E. Ragab ⁸

¹ Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah 47040, Pakistan

² Department of Computer Science, University of Education Lahore, Jauharabad Campus, Lahore 54770, Pakistan

³ Department of Applied Data Science, Noroff University College, 4612 Kristiansand, Norway

⁴ Department of Electrical and Computer Engineering, Lebanese American University, Byblos P.O. Box 13-5053, Lebanon

⁵ Artificial Intelligence Research Center (AIRC), College of Engineering and Information Technology, Ajman University, Ajman 346, United Arab Emirates

⁶ MEU Research Unit, Middle East University, Amman 11831, Jordan

⁷ Centre for Smart Systems, AI and Cybersecurity, Staffordshire University, Stoke-on-Trent ST4 2DE, UK

⁸ Industrial Engineering Department, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia

* Correspondence: skadry@gmail.com

Abstract: Esophagitis, cancerous growths, bleeding, and ulcers are typical symptoms of gastrointestinal disorders, which account for a significant portion of human mortality. For both patients and doctors, traditional diagnostic methods can be exhausting. The major aim of this research is to propose a hybrid method that can accurately diagnose the gastrointestinal tract abnormalities and promote early treatment that will be helpful in reducing the death cases. The major phases of the proposed method are: Dataset Augmentation, Preprocessing, Features Engineering (Features Extraction, Fusion, Optimization), and Classification. Image enhancement is performed using hybrid contrast stretching algorithms. Deep Learning features are extracted through transfer learning from the ResNet18 model and the proposed XcepNet23 model. The obtained deep features are ensembled with the texture features. The ensemble feature vector is optimized using the Binary Dragonfly algorithm (BDA), Moth–Flame Optimization (MFO) algorithm, and Particle Swarm Optimization (PSO) algorithm. In this research, two datasets (Hybrid dataset and Kvasir-V1 dataset) consisting of five and eight classes, respectively, are utilized. Compared to the most recent methods, the accuracy achieved by the proposed method on both datasets was superior. The Q_SVM's accuracies on the Hybrid dataset, which was 100%, and the Kvasir-V1 dataset, which was 99.24%, were both promising.

Keywords: gastrointestinal tract disease; deep learning; features optimization; stomach diseases; features fusion; wireless capsule endoscopy; stomach cancer



Citation: Naz, J.; Sharif, M.I.; Sharif, M.I.; Kadry, S.; Rauf, H.T.; Ragab, A.E. A Comparative Analysis of Optimization Algorithms for Gastrointestinal Abnormalities Recognition and Classification Based on Ensemble XcepNet23 and ResNet18 Features. *Biomedicines* **2023**, *11*, 1723. <https://doi.org/10.3390/biomedicines11061723>

Academic Editor: Nelson Yee

Received: 3 March 2023

Revised: 23 May 2023

Accepted: 9 June 2023

Published: 15 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Numerous researchers are recommending their methods for the precise identification and classification of gastrointestinal anomalies, skin lesions, and brain tumors. In the area of computer vision and medical image processing, this is a growing field of study. These research studies are usually based on the dataset obtained through imaging technologies such as computed tomography (CT), wireless capsule endoscopy (WCE), and magnetic resonance imaging (MRI). Gastric abnormalities incorporate esophagitis, bleeding, ulcer, and polyps. The most common gastrointestinal abnormalities that humans suffer from are bleeding, polyps, and ulcers [1]. These stomach abnormalities have turned into a main

source of mortalities in people [2]. Around the world, stomach disease is the third-most significant reason for death among all malignant deaths [3]. Due to the recent environmental changes and ecological contamination in human dietary propensities, gastrointestinal system sicknesses have become progressively serious dangers to human wellbeing. The detection of Gastrointestinal Tract (GIT) tumors at the early stages will prevent many other GIT diseases from being more severe [4]. Since the last two decades' rapid growth of technology, the utilization of machine learning, computer vision, and deep learning algorithms resulted in the extraordinary performance of several domains such as plants disease detection, surveillance, medical imaging for infections, and severe disease detection such as tumors [5]. Esophageal and gastric cancers are both common and deadly. Patients present most often after disease progression, and survival is therefore poor. Due to demographic variability and recent changes in disease incidence, much emphasis has been placed on studying risk factors for both esophageal and gastric cancers. Esophageal cancer (EsC), including squamous cell carcinoma (SCC) and adenocarcinoma, is considered as a serious malignancy with respect to prognosis and a fatal outcome in the great majority of cases. Esophageal carcinoma affects more than 450,000 people worldwide, and the incidence is rapidly increasing [6,7]. Gastric cancer accounts for 783,000 deaths each year, making it the third-most deadly cancer among males worldwide. In total, 8.3% of all cancer deaths are attributable to gastric cancer. The cumulative risk of death from gastric cancer, from birth to age 74, is 1.36% for males and 0.57% for females. Esophageal cancer can cause bleeding, and according to a publication in the American Journal of Gastroenterology, individuals with esophagitis have a significantly increased risk of developing esophageal cancer [8].

WCE is a noninvasive diagnostic method used for the careful examination of the bowel [9]. It is a painless examination as compared with existing traditional techniques [10,11]. WCE images play a very decisive role in the evaluation of stomach bleeding such as infections in the gastric tract (GI) and ulcers. A well-known technique, video capsule endoscopy (VCE), is also used for it [11,12]. When WCE is passed down the throat, it captures colored images for a medium time of 8 h that are received by a data recording device. Then, expert physicians analyze these images and make decisions about the disease [13]. It is a fact that standard endoscopic techniques diagnose the disease correctly, but they have the main disadvantage of causing discomfort and pain. To resolve these issues, the WCE modality is used. Many researchers have utilized WCE images for the evaluation of their proposed methods. Image classification is the most challenging task, especially in medical imaging. Doctors must spend a lot of time and effort categorizing and identifying ulcers, but doing so is crucial. To solve this problem, however, computerized methods have been created by many researchers. These techniques mainly rely on (a) familiarity with diverse ulcers, bleeding, and polyp images, (b) features such as color and texture, and (c) accuracy and execution time calculation. Researchers have recently developed a number of methods for automatically identifying diseased regions, especially focused on the color, texture, shape, and deep learning features. Due to difficulties including low-quality dataset pictures, color similarity between infected and healthy parts, differences in the texture and form of the infected area, and the necessity for feature engineering, detecting and classifying gastrointestinal disorders is a challenging endeavor [14].

The primary aim of this research is to propose a method that will recognize gastrointestinal abnormalities accurately with reduced computational time and improved results in terms of accuracy. The proposed method is based on the two deep learning models. There exist several challenges in this domain, such as images with low quality, noisy images, the selection of the appropriate method for quality enhancement, and suitable features extraction. Improving the quality of the dataset such that regions of interest become more highlighted and subtle information becomes prominent, strong feature selection, variation in the scalability of input data, and variation in texture information are also challenges. A gastrointestinal abnormality recognition and classification strategy is presented to tackle these limitations. The main steps of our proposed method are: (a) image pre-processing; (b) features engineering; (c) classification. The major contributions are:

- Contrast enhancement is one of the main contributions of this research work. In the pre-processing step, hybrid contrast enhancement methods are used to improve lesion contrast. The major steps performed in the pre-processing phase are 3D-box filtering, 3D-median filtering, HSI color transformation, and channels extraction. Contrast Limited Adaptive Histogram Equalization (CLAHE) is employed on the extracted channels and gives the output H_CLAHE, S_CLAHE, and I_CLAHE; in the next step, the saturation Weight-Map is applied on the S_CLAHE. In the last step of the pre-processing phase, un-sharp masking is employed on the concatenated output.
- A CNN model is designed from scratch with 23 layers, named XcepNet23. The proposed CNN model is pretrained on the CIPHER-100 dataset. Features engineering is performed on the extracted deep features.
- A hybrid features engineering methodology is proposed. Features vectors are obtained using the fine-tuned ResNet18 model, the proposed XcepNet23, and LBP features. The strong feature vector is created by combining the texture features and the extracted features from the two CNN models. The BDA, MFO, and PSO algorithms are used to optimize the fused feature vector.

The main objective is to develop a reliable and efficient model that can correctly identify and categorize digestive diseases. In comparison to current techniques, the research intends to obtain improved diagnostic accuracy, facilitating early identification and efficient treatment. The objective is to identify and categorize a wide range of digestive diseases by utilizing the strength of CNN and ML algorithms. This seeks to give professionals a thorough diagnostic tool for precisely identifying diseases. The rest of the paper is arranged as follows: The literature review is in Section 2, the concise narrative of the proposed methodology is in Section 3, and the quantitative experimental results are discussed in Section 4. Section 6 presents the paper's conclusion.

2. Related Work

The automated detection of disease is an active area of research. Researchers have introduced numerous automated computer-aided diagnosis methods to help physicians. Most of the methods are suggested for malignant disease detection, including brain tumor segmentation and classification [15,16], glaucoma detection [17], lung cancer detection [18], and so on. The accurate detection of the infected region, such as ulcers and bleeding, is a most challenging task. There exist different methods that the researchers use for image enhancement based on gamma correction [19], image colorization [20], a geometric filter [21], the OTSU Threshold [22] and discrete Fourier-transform. A large memory and normally more than 8 h are required for the WCE examination. A real-time analysis algorithm is required to correctly analyze abnormal tissues or infected regions.

An algorithm is developed for automatic bleeding region detection [23] that mainly focuses on bleeding spot detection. It consists of the statistical analysis and shape analysis of the region of interest (ROI). This algorithm was tested on 30 different cases of capsule endoscopy and achieved 97% and 99% accuracy of specificity and sensitivity, respectively. Yuan et al. [24] came up with an approach for the detection of ulcers, polyps, and bleeding using the K-means clustering algorithm. They utilized WCE images and achieved an 88.61% average performance. The authors presented a procedure for bleeding region recognition from WCE imageries. In this method, pixels are assembled robustly based on the intensity and area through superpixel division instead of processing each pixel separately or through a uniform division of pixels [25]. A novel method is used in [26] for distinguishing healthy and bleeding regions. Bleeding is a major symptom of a disease. Different color models are used for the detection of bleeding, and experiments are performed by using local binary pattern (LBP). Experiments show that by using CIE XYZ, an accuracy of 96.38% is achieved for the KNN classifier. The identification of problematic images has been an obstacle for specialists to deal with. An integrated saliency measure and the Bag of Features method are suggested to tackle this problem [27]. The proposed technique provides good classification and characteristics of polyps from WCE images. In [28], diverse measurable parameters are

thought to accurately detect bleeding images such as the maxima, mean, minima, median, mode, variance, kurtosis, median, and skewness. In [29], images of ulcers and healthy people are categorized using the extraction of texture feature extraction. The contour let transform and log Gabor filter are used for texture features extraction. SVM receives the extracted texture features, which have the highest accuracy of 94.16%. Yuan et al. [30] suggested a saliency-based practice for the detection of ulcers. The proposed method comprises two main stages. Sample images of the ulcer are selected in the first step, whereas in step two, the multilevel super pixel method is used to segment infected regions that are joined in a single map. In [31], the authors proposed an approach for separating healthy and unhealthy pixels through different color spaces. RGB, HSV, YCbCr, etc. shading spaces are utilized for the extraction of texture features. Reduced FV is supplied to the SVM classifier for the classification of reduced features. They have shown a 97.89% classification accuracy. Reed T. Sutton et al. [32] carried out research in which they compared the performance of several deep learning algorithms. The authors utilized the HyperKvasir dataset and obtained 8000 images from the dataset. From the comparison of the CNN model, it is concluded that the DenseNet121 model achieved a maximum accuracy of 87.50%. A method for identifying stomach diseases was proposed by Nayar et al. [33]. The authors combined the CNN features with the improved Genetic Algorithm (GA). The feature extraction in this study is carried out with AlexNet. The computational time required for this method was 211.90 s, and the accuracy was 99.8%. Guanghua Zhang et al. [34] proposed a method for digestive tract tumor detection. In our previous study [35], a hybrid method was proposed for the classification of gastrointestinal diseases. The major steps involved in this methodology are pre-processing, texture features extraction, CNN features extraction, feature fusion, and classification. The highest classification accuracy achieved on the KVASIR dataset was 99.3%. The literature review depicts the pre-processing and features extraction steps used by most researchers for disease classification and detection. Moreover, some researchers have utilized a combination of different types of features for better classification results. As the prediction of ulcers, health, and bleeding from WCE images is greatly influenced by features that are to be extracted, therefore, in this research, our focus will be on the pre-processing step and features engineering phase.

A crucial phase that has a significant impact on prediction performance is improving the quality of the dataset. A hybrid strategy for enhancing the image quality is proposed in this study. The feature's engineering phase is likewise basic in our proposed technique, wherein two kinds of features are extracted. These features are obtained utilizing CNN models and the Local Binary Pattern (LBP) algorithm. The CNN features are extracted using the proposed CNN model (XcepNet23) and the ResNet18 model. In the final phase, machine learning algorithms are utilized for classification. The following section provides a comprehensive explanation of the proposed method.

3. Proposed Methodology

The automatic detection of gastrointestinal diseases is a trending area of research in computer vision and image processing. Due to the identical color of cells in normal and infected WCE images, the classification of images is also an important task that is highly sensitive. The proposed method's in-depth description is included in this section. This research proposes a method consisting of hybrid image processing, machine learning, and deep learning algorithms for the recognition of gastrointestinal abnormalities. This work is primarily based on five main phases, which are: (a) image augmentation; (b) hybrid pre-processing technique; (c) features extraction; (d) feature fusion and optimization; and (e) classification. The following sections provide a comprehensive explanation of these phases. The model that represents the proposed method is given in Figure 1.

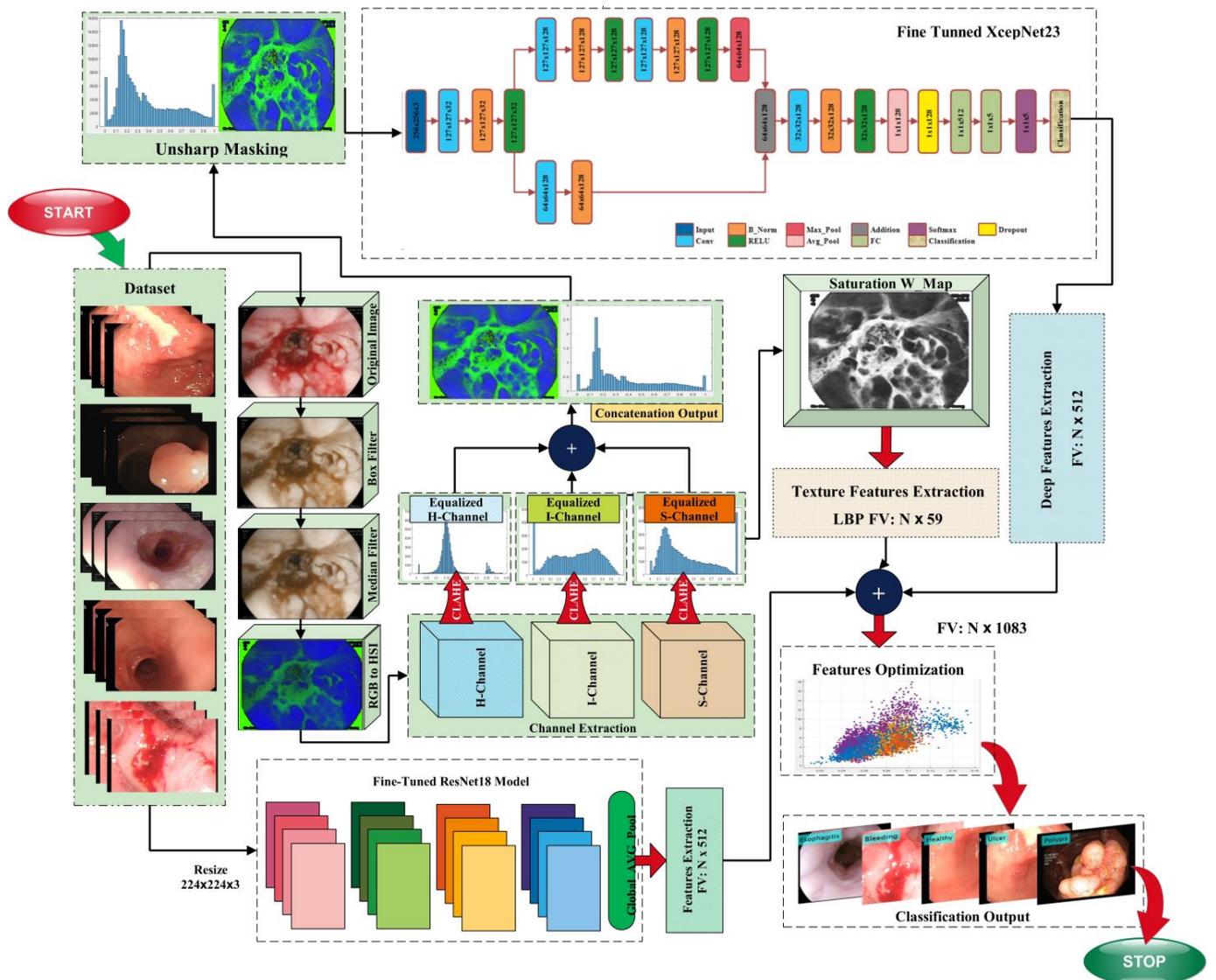


Figure 1. Proposed model block diagram.

3.1. Preprocessing

Preprocessing is the first phase of our method; the contrast of the image is improved to enhance the chromatic quality of the infected region. The pre-processing step has a great impact on different domains such as computer vision, medical imaging, biometrics, and surveillance. This step is performed in medical imaging due to the presence of numerous challenges such as the low contrast of the images, the presence of blur, noise artifacts, variations in illumination, color variations, and lightning effects. Image enhancement is performed to extract the most salient features and obtain accurate classification results. In this phase, different methods are utilized for contrast enhancement. The pre-processing step is further subdivided into major sub-steps, which are: data augmentation, 3D-box filtering, 3D-median filtering, RGB to HSI color transformation, channels extraction (H, S, I), CLAHE is employed on the extracted channels, which results in H_CLAHE, S_CLAHE, and I_CLAHE, a saturation weight map is employed on the S_CLAHE, and feature engineering is employed on the ensemble feature vector.

3.1.1. Data Augmentation

The availability of large datasets is still a challenge for researchers. For the training phase, deep learning models require a sizable quantity of data. Overfitting is a well-known

issue that affects all researchers and is brought on using insufficient or tiny datasets. In this study, we used the data augmentation technique to expand the dataset size and balance the dataset. The research reveals that there are several image augmentation techniques, including image-flipping, mirrored images, and image rotation. In this paper, image rotation is utilized for the lossless augmentation of the dataset. We have rotated the original image in three angles, which are: 90°—right, 180°—right, and 270°—right (or 90°—left). Some sample output images of data augmentation are given in Figure 2.

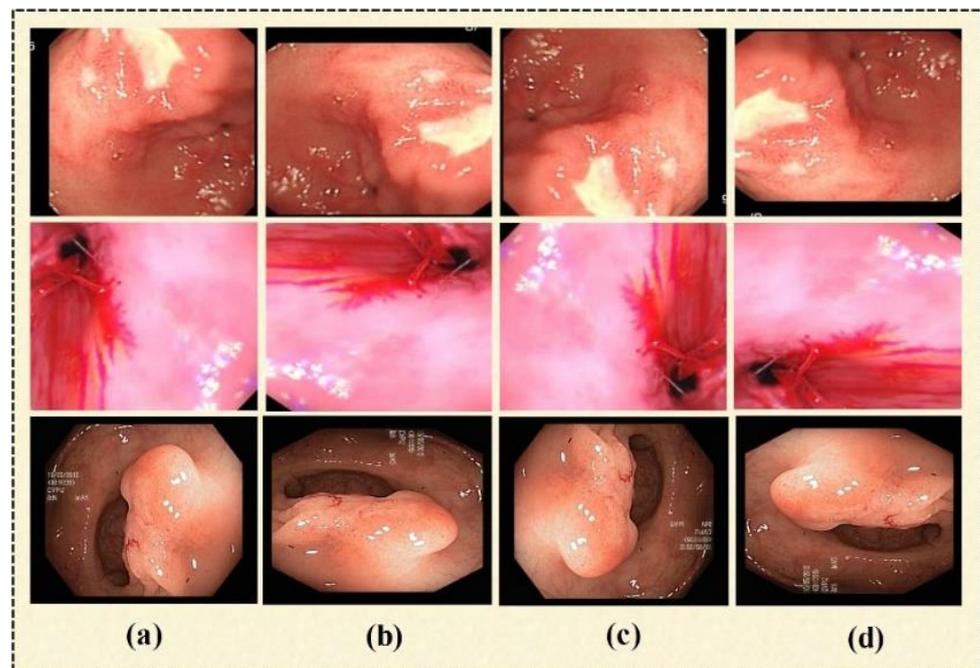


Figure 2. Sample images. (a) Original image, (b) 90°—right, (c) 180°—right, (d) 90°—left.

3.1.2. D-Box Filtering

In this step of image preprocessing, a 3D-box filter [12] is employed for the initial image, as shown in Figure 3a. The output image of this step is shown in Figure 3b. It is used to smooth the image based on assigning an equal mask to the neighborhood pixels. It has three steps. In the first step, an RGB image $I(s,d)$ is converted into three channels. A box filter (1's pixel) is generated, which a size of 3×3 . A mathematical representation of the box filter is given in Equation (1).

$$\text{Mask}(h,v) = \sum_{j=1}^3 \sum_{k=1}^3 [H_j][V_k] \quad (1)$$

where $j \in H$, $k \in V$, and h and v represent the rows and columns of the produced mask filter. Here, the mask size is $3 \times 3 = 9$, which is separately applied on all channels, as depicted in Equations (2)–(4).

$$\mathcal{R}_{\text{mask}}(s,d) = \text{Mask}(h,v)[\mathcal{R}(s,d)] \quad (2)$$

$$\mathcal{G}_{\text{mask}}(s,d) = \text{Mask}(h,v)[\mathcal{G}(s,d)] \quad (3)$$

$$\mathcal{B}_{\text{mask}}(s,d) = \text{Mask}(h,v)[\mathcal{B}(s,d)] \quad (4)$$

where $\mathcal{R}(s,d)$ is the red channel and $\mathcal{G}(s,d)$ and $\mathcal{B}(s,d)$ represent the green and blue channels, respectively. $\mathcal{R}_{\text{mask}}(s,d)$ represents the kernel filter for the red channel, $\mathcal{G}_{\text{mask}}(s,d)$ is the kernel filter for the red channel, and $\mathcal{B}_{\text{mask}}(s,d)$ is the kernel filter for the blue channel. Equations (5)–(7) for channel extraction are as follows.

$$\mathcal{R}(s, d) = \frac{\text{red}}{I(s, d)} \quad (5)$$

$$\mathcal{G}(s, d) = \frac{\text{green}}{I(s, d)} \quad (6)$$

$$\mathcal{B}(s, d) = \frac{\text{blue}}{I(s, d)} \quad (7)$$

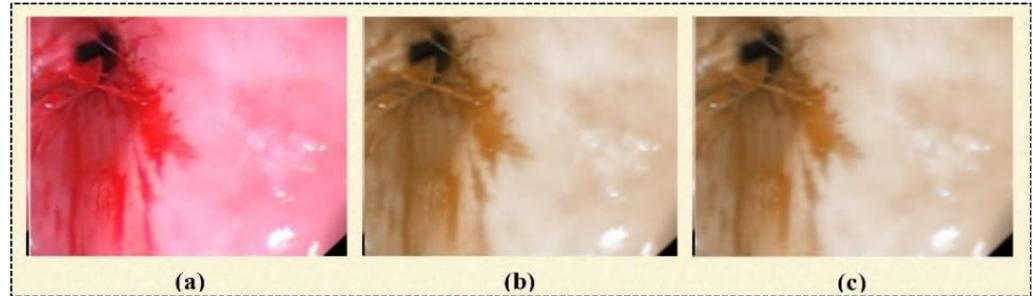


Figure 3. (a) Original image, (b) box-filter output image, (c) median-filter output image.

Updated pixel values and kernel filters are utilized to perform convolution operations. A new matrix is produced by using a null matrix with a size of 256×256 and the above-generated values of $h = 3$, $v = 3$, $h_1 = 256$, and $v_1 = 256$. The resultant new matrix is shown below in Equation (8).

$$F_c(s, d) = \sum_{j=1}^{h_M} \sum_{k=1}^{v_N} [Z_{j,k}(s, d) + \text{Mask}(h, v)] \quad (8)$$

where $j \in h_u$, $k \in h_c$, $h_u = 3$, $v_u = 3$, $h_M = h_u + h_1 - 1$, and $v_N = v_u + v_1 - 1$. The symbols h_u and v_u represent updated values of rows and columns iterated up to 256 times. 2D images are produced by using $F_c(s, d)$ for each channel in the last step. Concatenating them results in a new, improved picture that amplifies the contrast between the diseased area and the surrounding area, as explained in Equations (9) and (10), respectively.

$$B_F(s, d) = \sum_{j=1}^M \sum_{k=1}^N I_{j,k} \quad (9)$$

$$F_n(s, d) = \sum_j B_Fj(sd) \quad (10)$$

where $F_n(s, d)$ represents the result of the 3D-box filter, as given in Figure 3b.

3.1.3. D-Median Filtering

It is applied to the output image produced by the 3D-box filter. Noise issues are resolved through this filter [36]. The resultant image is shown in Figure 3c.

3.1.4. HSI Color Transformation

The additive system, which can be obtained by combining all three channels—red, green, and blue—is the foundation of the RGB color model. The numbers that represent these colors are expressed as the set of three numeric values representing the intensity of three channels ranging from 0 to 255. The lowest value represents the black color (0, 0, 0), while the highest value is the representation of the white color (255, 255, 255). Another color model based on three parameters is the HSI model. In HSI color transformation, colors are encoded according to Hue, Saturation, and Intensity. Along with the RGB color model, this color system is used as an alternative to RGB color space in some color monitors [37]. RGB to HSI color transformation is applied to the outcome from the previous step. In this color wheel, the hue of the HIS transformation of intensity or color is represented by the

angle measure. The second channel of the HSI color model is saturation. A color's intensity or purity, especially in relation to its brightness, is referred to as saturation. It shows how much a color has been diluted with white or grey. The highly saturated color seems to be pure and vivid whereas, the de-saturated image is washed out. Low saturation indicates that there is dilution with white light, whereas high saturation means the color is pure and free from dilution. The third component is the intensity (brightness), which depends on the saturation and hue component. The mathematical equations for the conversion of RGB to the HSI color space [38] are given in Equations (11)–(14):

$$\mathcal{I} = (r + g + \ell) / 3 \quad (11)$$

$$\mathcal{H} = \cos^{-1} \left(\frac{0.5[(r - g) + (r - \ell)]}{\sqrt{(r - g)^2 + (r - \ell)(g - \ell)}} \right) \quad (12)$$

$$\text{if } \ell > g \text{ then } \mathcal{H} = 360^\circ - \mathcal{H} \quad (13)$$

$$\mathcal{S} = 1 - \left(\frac{3}{r + g + \ell} \right) [\min(r, g, \ell)] \quad (14)$$

3.1.5. Contrast Limited Adaptive Histogram Equalization

Numerous researchers in image processing have used CLAHE to improve contrast by calculating distinct histograms for each image region; this algorithm improves contrast. These histograms are utilized in the next steps for the equal distribution of gray levels across the range of images. After the color transformation, in the next step, the H, S, and I channels are extracted, and CLAHE [39] is applied to all three extracted channels. After CLAHE, three output images are obtained, which are H_CLAHE, S_CLAHE, and I_CLAHE. The ensemble feature vector ($\phi_{Ensemble}^{CLAHE}$) is obtained by concatenating the H_CLAHE, S_CLAHE, and I_CLAHE. After the CLAHE step, the Chromatic weight map (C_W_MAP) is employed on the S_CLAHE, and an improved image is obtained ($S_CLAHE_{Improved}$) as an output of this step. The other two CLAHE outputs remain unchanged. The $S_CLAHE_{Improved}$ is further utilized for the texture features extraction. The $\phi_{Ensemble}^{CLAHE}$ is sharpened to improve the quality. In the next phase, the $\phi_{Ensemble}^{SHARP}$ image is obtained, which is further utilized for the proposed CNN model for the deep features extraction.

3.2. Feature Extraction

Feature Extraction and optimal feature selection are significant phases that are necessary to achieve the accurate detection and classification of GIT abnormalities. In this research, we have utilized the Local Binary Pattern for the texture features extraction, and two CNN model are utilized for the deep features extraction. In Figure 4, there is an overview of the phases from preprocessing features extraction, deep features extraction, and features fusion. Whereas in the preprocessing phase we have performed several steps such as data augmentation, box-filtering, median filtering, and RGB-to-HSI conversion, in the preprocessing block of Figure 4, step-i represents the extracted Equalized S-channel output of the preprocessing phase, and Step-n represents the final output image of the preprocessing phase. The equalized S-channel image is utilized for the LBP features extraction that is later ensembled with the extracted deep learning features. A block diagram representing the proposed features engineering process is given in Figure 4.

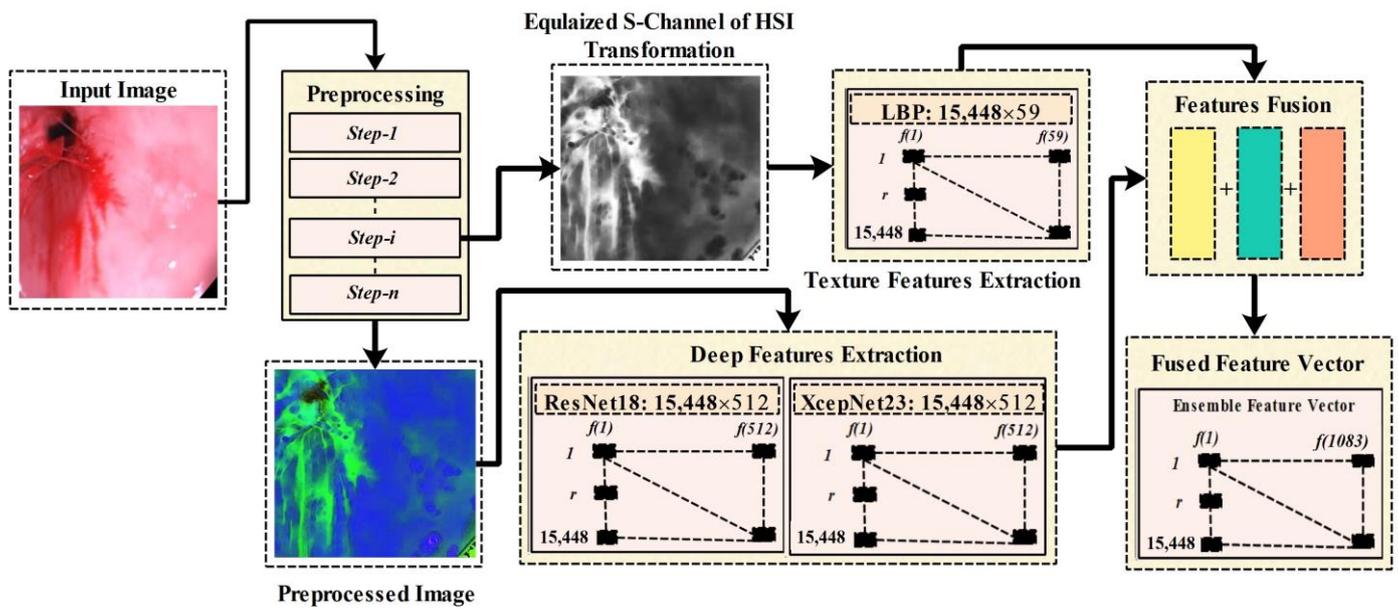


Figure 4. Proposed features engineering process.

3.2.1. Local Binary Pattern

LBP [40] is an important method utilized by many researchers for the extraction of texture information. It takes a grayscale image as the input after that variance and mean are calculated for all of the intensities (pixels) of the image. In this research project, we have utilized the S_CLAHE for the LBP features extraction. This algorithm returned the feature vector of dimension $\mathcal{N} \times 59$. The mathematical formulation of LBP is given below in Equations (15) and (16):

$$LBP(Q, \mathcal{R}) = \sum_{Q=0}^{Q-1} \mathfrak{Y}(\mathcal{E}_Q - \mathcal{E}_{\mathcal{R}})2^Q \tag{15}$$

$$\mathfrak{Y}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

In this formulation, the neighborhood intensities are given by \mathcal{Q} , the radius is given by \mathcal{R} , $\mathcal{E}_{\mathcal{R}}$ represents the current pixel intensity, and \mathcal{E}_Q represents the neighboring pixel intensity. $\mathfrak{Y}(x)$ refers to the thresholding function.

3.2.2. Deep Learning Features

Inspired by the usefulness of deep learning in the surveillance and agriculture domain, several researchers have employed deep learning models for disease recognition. Deep learning models have resulted in a great breakthrough in terms of disease detection and classification performance. A CNN model consists of multiple layers comprising the input, convolution, batch normalization, RELU, pooling, softmax, and classification layers. Each layer is responsible for performing a specific task. A detailed description of the CNN models utilized in this research is given in the following section:

ResNet18

There are different versions of Residual Network (ResNetXX) architecture [41], where “XX” specifies the total number of layers. For the deep feature extraction in this study, the ResNet18 version was used. There are 72 layers in this architecture, with 18 of them representing the model’s 18 deep layers. The ResNet18 model is originally pre-trained on the thousand categories of the ImageNet dataset, consisting of the ResNet18 model

3.2.3. Features Optimization

In the literature, we have observed that authors have utilized feature fusion to obtain a strong resultant set of features. There may be some redundant features that reduce the performance of the method. In the last decade, methods based on metaheuristic algorithms have been viewed as the most proficient and solid improvement strategies. These algorithms have been broadly utilized for the improvement of the performance of real-world issues. In this research, we have utilized feature optimization algorithms to extract the optimal features among the total extracted features. Optimization algorithms help to get rid of redundant and non-optimal features. Many algorithms can be used for selecting the best features. It is observed in the literature that authors have utilized BDA [2], MFO in combination with the Crow Search algorithm (CSA) [42], PSO and CSA [43], Enhanced Crow Search and Differential Evaluation, and Grasshopper [44]. Inspired by these studies, in this paper, we used three algorithms inspired by nature to optimize features, which are BDA, MFO, and PSO. The following sections briefly overview the various optimization techniques under consideration. We have one feature vector after the features fusion step denoted by ϕ^{Fused} , having the dimension $N \times 1083$. Optimization algorithms are applied on the ϕ^{Fused} to obtain the optimal set of features. The obtained feature vector is optimized using BDA, MFO, and PSO. The mathematical formulation of optimization algorithms is as follow:

The BDA [45] mimics the swarming dragonflies (DF) pattern. The fact-finding and shady systems of DA [46] are demonstrated by the association of dragonflies in keeping away from the foe and finding the source of food. The dragonfly algorithm mathematical formulation is given in the following Equations (17)–(25). Separation is formulated as:

$$\mathfrak{S}_i = -\sum_{i=1}^N \mathfrak{x} - \mathfrak{x}_i \quad (17)$$

where \mathfrak{x} specifies a DF position in a space that is M-dimensional. N and \mathfrak{x}_i represent the number of neighbor individuals and the position of the neighbor individual, respectively.

Alignment permits speed matching of the individual in a sub-swarm/swarm. The mathematical formulation of Alignment is as follows:

$$\mathfrak{A}_i = \frac{\sum_{i=1}^N \mathfrak{V}_i}{N} \quad (18)$$

where N represents the neighbor individuals and \mathfrak{V}_i represents the neighbor's individual velocity. Cohesion alludes to the deviation of the ongoing individual toward the focal point of the mass of the neighbor individual.

$$\mathfrak{C}_i = \frac{\sum_{i=1}^N \mathfrak{x}_i}{N} - \mathfrak{x} \quad (19)$$

where \mathfrak{x}_i specifies the individual neighbor position.

Attraction is also an important behavior; it specifies that the source of food should be the attraction for the individuals. Mathematically, the attraction behavior is formulated as:

$$\mathfrak{L}_\varphi = \mathfrak{x}_\varphi - \mathfrak{x} \quad (20)$$

where \mathfrak{x}_φ indicates the source food's position.

Distraction indicates the situation in which an individual should not get closer to the enemy. The individual should get away from predators. The mathematical formulation of distraction is as follows:

$$\mathfrak{E}_i = \mathfrak{x}_\mathfrak{E} - \mathfrak{x} \quad (21)$$

where $\mathfrak{x}_\mathfrak{E}$ represents the location of the enemy.

These five behaviors control the DFs' movement in DA. The step vector is calculated to update the location of each dragonfly:

$$\Delta \mathfrak{X}_i(\mathcal{T} + 1) = (s\mathfrak{S}_i + \alpha \mathfrak{A}_i + c\mathfrak{C}_i + f\mathfrak{L}_{\mathcal{P}} + e\mathfrak{E}_i) + w\Delta \mathfrak{X}_i(\mathcal{T}) \tag{22}$$

where \mathcal{T} signifies the current iteration, s is the separation weight, α is the alignment weight, c is the weight of cohesion, f is the food weight, e is the weight of the predator, and w is the weight of inertia. In the initial algorithm for dragonflies, the locations of the DFs are restructured using the following equation:

$$\Delta \mathfrak{X}_i(\mathcal{T} + 1) = \mathfrak{X}_i(\mathcal{T}) + \Delta \mathfrak{X}_i(\mathcal{T} + 1) \tag{23}$$

The locations vectors of BDA are updated as follows:

$$\mathfrak{X}_i^k(\mathcal{T} + 1) = \begin{cases} 1 - \mathfrak{X}_i^k(\mathcal{T}) & \text{Rand} < \mathfrak{T}\ell(\Delta \mathfrak{X}_i^k(\mathcal{T} + 1)) \\ \mathfrak{X}_i^k(\mathcal{T}) & \text{Rand} \geq \mathfrak{T}\ell(\Delta \mathfrak{X}_i^k(\mathcal{T} + 1)) \end{cases} \tag{24}$$

$$\mathfrak{T}\ell(\Delta \mathfrak{X}) = \left| \frac{\Delta \mathfrak{X}}{\sqrt{\Delta \mathfrak{X}^2 + 1}} \right| \tag{25}$$

where Rand , $\Delta \mathfrak{X}$, \mathfrak{X}_i^k , and $\mathfrak{T}\ell$ represent the number that is randomly generated (range: 0–1), the step vector, the k th position of the i th dragonfly, and the transfer function, respectively. The ensemble feature vector ϕ^{Fused} is processed using the BDA algorithm and obtains the optimized feature vector with the dimensions $\mathcal{N} \times 384$.

In this research, we have also utilized the Moth–Flame Optimization algorithm [47] for feature optimization. There are three major steps that the moths follow: population creation, position updating, and updating the final amount of the flame. Here, the Moths population and Fitness function value can be mathematically expressed as given in Equations (26)–(34):

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,n} \\ w_{2,1} & w_{1,2} & \dots & w_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{z,1} & w_{z,2} & \dots & w_{z,n} \end{bmatrix} \tag{26}$$

$$OW = \begin{bmatrix} OW_1 \\ OW_2 \\ \vdots \\ OW_z \end{bmatrix} \tag{27}$$

$$\mathfrak{F} = \begin{bmatrix} \mathfrak{F}_{1,1} & \mathfrak{F}_{1,2} & \dots & \mathfrak{F}_{1,n} \\ \mathfrak{F}_{2,1} & \mathfrak{F}_{1,2} & \dots & \mathfrak{F}_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{F}_{z,1} & \mathfrak{F}_{z,2} & \dots & \mathfrak{F}_{z,n} \end{bmatrix} \tag{28}$$

$$O\mathfrak{F} = \begin{bmatrix} O\mathfrak{F}_1 \\ O\mathfrak{F}_2 \\ \vdots \\ O\mathfrak{F}_z \end{bmatrix} \tag{29}$$

where the matrix W represents the moth’s population and the flame is represented by the F , n , and z denoting the number of dimensions and the number of moths. After that, the moths’ positions are altered to obtain the best global solutions. The following function is chosen for the optimization challenge’s best global solution:

$$MFO = (P, Q, S) \quad (30)$$

where P is a function, P generates the fitness function corresponding to the randomly generated population $P : \emptyset \rightarrow \{W, OW\}$, and the Q function is the main function. Q is responsible for the moth movement around the search space $Q : W \rightarrow W$, and the S function returns false or true based on the termination criteria S, which is formulated as $S : \{true, false\}$. The mathematical formulation of the moth logarithmic spiral for the updating mechanism is given as follows:

$$W_i = S(W_i, \mathfrak{F}_j) \quad (31)$$

$$S(W_i, \mathfrak{F}_j) = \mathcal{D}_i \cdot e^{cr} \cdot \cos(2\pi r) + \mathfrak{F}_j \quad (32)$$

$$\mathcal{D}_i = |\mathfrak{F}_j - W_i| \quad (33)$$

$$Selected_{\mathfrak{F}_{lame}} = round\left(\mathfrak{M} - \mathfrak{k} * \frac{\mathfrak{M} - 1}{\mathbb{T}}\right) \quad (34)$$

where \mathcal{D}_i indicates the distance between the j th flame and the i th moth, c is a constant that defines the logarithmic spiral shape, and r is a random no. between the range -1 and 1 . Equation (34) \mathfrak{M} denotes the max number of flames, the total number of iterations is denoted by \mathbb{T} , and \mathfrak{k} represents the current iteration. $Selected_{\mathfrak{F}_{lame}}$ is passed to the Equations (26)–(29), and the fitness function is calculated. The MFO technique that was mentioned earlier is used to process the fused feature vector ϕ^{Fused} and obtain the optimized feature vector with the dimensions $\mathcal{N} \times 437$.

In this article, the third optimization algorithm that we have utilized is the PSO Algorithm [48]. PSO is a nature-inspired algorithm that is derived from the survival behavior of swarms. This algorithm's primary goal is optimization or picking the best solution from the available possible solutions. For a K -Dimensional search space, where we have k particles, the vector for the j th particle of the swarm will be represented as $\mathcal{P}_j = (p_{j1}, p_{j2}, \dots, p_{jK})$, whereas, the prior optimal position of the j th particle is $\ell_j = (\ell_{j1}, \ell_{j2}, \dots, \ell_{jK})$, which returns the optimal fitness value. Here, \mathcal{L} symbolizes the particle's lowest function value, and the velocity of the j th particle can be expressed as $\mathcal{V}_j = (v_{j1}, v_{j2}, \dots, v_{jK})$. The following mathematical formulation represents the particles' updating process:

$$v_{id}^{y+1} = \mathcal{J} \times v_{id}^y + c1 \times r() \times (\ell_{id}^y - p_{id}^y) + c2 \times R() \times (\ell_{gd}^y - p_{gd}^y) \quad (35)$$

$$p_{id} = p_{id} + v_{id} \quad (36)$$

where $c1$ and $c2$ denote the positive constant numbers of cognitive and social parameters, respectively. Here, $R()$ and $r()$ are the two functions that produce pseudorandom numbers within a certain range $[0, 1]$. This algorithm returned a feature vector of $\mathcal{N} \times 428$ dimensions.

4. Experimental Results and Analysis

The dataset consists of thousands of images which are unhealthy and healthy. Unhealthy images are further categorized into two classes: ulcers and bleeding. In this work, several machine-learning algorithms are used to classify the data. To evaluate the findings presented in this paper, each of the six classifiers listed below is taken into consideration: Fine Tree (F_Tree) [49], Coarse Tree (C_Tree) [50], Quadratic SVM (Q_SVM) [51], Fine Gaussian SVM (F_G_SVM) [52], Coarse KNN (C_KNN) [53], and Ensemble Subspace Discriminant (E_S_Disc) [54]. The classifier models include Fine Tree (F_Tree) and Coarse Tree (C_Tree), which are decision tree-based models optimized for fine and coarse-grained classification tasks, respectively. Quadratic SVM (Q_SVM) is a support vector machine

model that utilizes kernel functions to handle non-linearly separable data. Fine Gaussian SVM (F_G_SVM) employs a Gaussian kernel to model the probability density function of the input data. Coarse KNN (C_KNN) is a k-nearest neighbor model optimized for coarse-grained classification tasks. Ensemble Subspace Discriminant (E_S_Disc) is an ensemble model that combines multiple classifiers operating on different subspaces of high-dimensional data. All the experiments are performed on MATLAB 2020a, Core i5-7th Generation with 8 GB RAM.

4.1. Dataset

Two datasets are utilized to evaluate the proposed method in this research paper. The considered datasets include: (1) the KVASIR_V1 Dataset and (2) the Hybrid_Dataset. The Hybrid_Dataset utilized in this research work consists of five classes: Bleeding, Ulcer, Esophagitis, Polyps, and Healthy. The dataset for this research is collected from different sources. Ulcer, Healthy, and Bleeding images are acquired from Amna Liaqat et al. [12] (each class contains 3000 images). The Esophagitis images are taken from the Kvasir dataset, [55]. Using the data augmentation technique, the volume of data for esophagitis images is increased without reducing the features in the original images. The fifth class of our dataset is polyps. Polyps' images are collected from two datasets named the Kvasir and CVC datasets. The proposed method is evaluated using a total of 15,448 images. The second dataset considered for this research work is the KVASIR_V1 dataset. This dataset consists of eight classes, and the total number of images in this dataset is 4000 [55]. We have used all classes of this dataset. Some sample images are given in Figure 7.

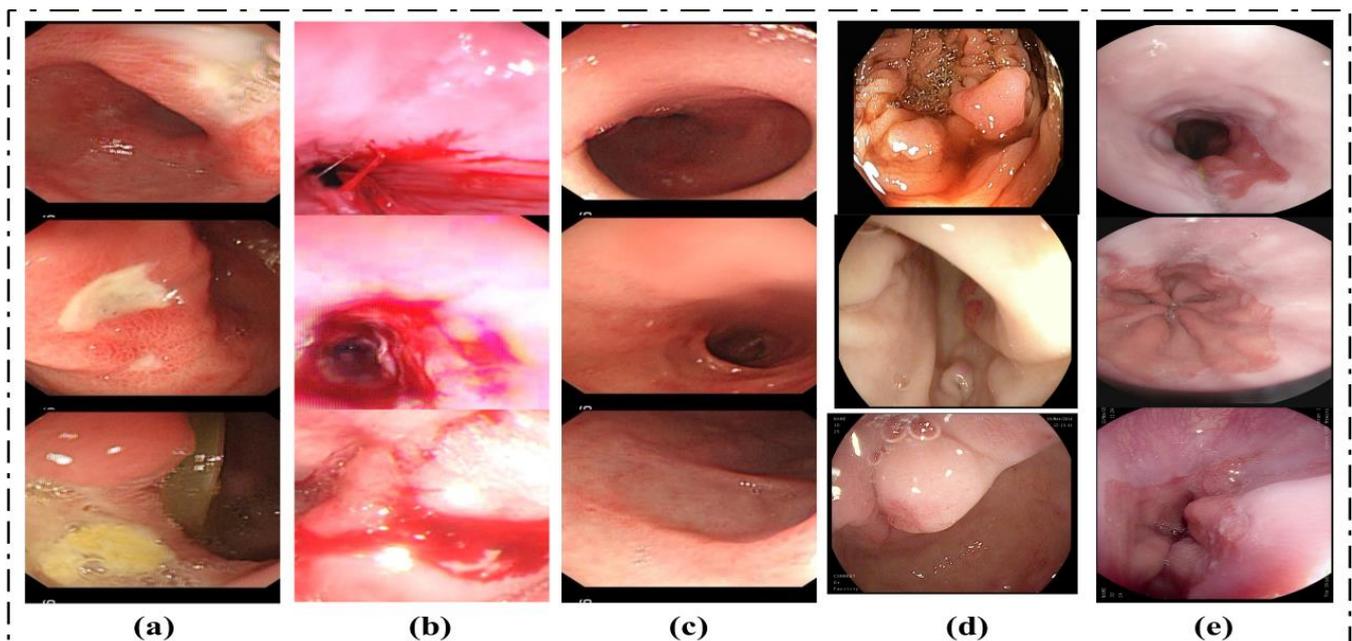


Figure 7. Sample dataset images: (a) Ulcer, (b) Bleeding, (c) Healthy, (d) Polyps, (e) Esophagitis.

4.2. Performance Measures

In this research, the execution of the suggested classification method is assessed on different performance procedures, which are the execution time, precision (PRE), accuracy (ACC), F1-Score (F1), specificity (SPE), sensitivity (SEN), Cohen's Kappa score, and Matthews correlation coefficient (MCC). Mathematically, these performance measures are given from Equations (37)–(43). In this research, we have utilized five-fold cross-validation to assess all of the quantitative results.

$$\text{SEN} = \frac{\mathcal{T}\mathcal{P}\mathcal{V}}{\mathcal{T}\mathcal{P}\mathcal{V} + \mathcal{F}\mathcal{N}\mathcal{V}} \quad (37)$$

$$\text{SPE} = \frac{\mathcal{T}\mathcal{N}\mathcal{V}}{\mathcal{F}\mathcal{P}\mathcal{V} + \mathcal{T}\mathcal{N}\mathcal{V}} \quad (38)$$

$$\text{PRE} = \frac{\mathcal{T}\mathcal{P}\mathcal{V}}{\mathcal{T}\mathcal{P}\mathcal{V} + \mathcal{F}\mathcal{P}\mathcal{V}} \quad (39)$$

$$\text{F1} = 2 * \left(\frac{\text{SEN} * \text{PRE}}{\text{SEN} + \text{PRE}} \right) \quad (40)$$

$$\text{ACC} = \frac{\mathcal{T}\mathcal{P}\mathcal{V} + \mathcal{T}\mathcal{N}\mathcal{V}}{\mathcal{T}\mathcal{P}\mathcal{V} + \mathcal{T}\mathcal{N}\mathcal{V} + \mathcal{F}\mathcal{P}\mathcal{V} + \mathcal{F}\mathcal{N}\mathcal{V}} \quad (41)$$

$$\text{MCC} = \frac{\mathcal{T}\mathcal{P}\mathcal{V} \times \mathcal{T}\mathcal{N}\mathcal{V} - \mathcal{F}\mathcal{P}\mathcal{V} \times \mathcal{F}\mathcal{N}\mathcal{V}}{\sqrt{(\mathcal{T}\mathcal{P}\mathcal{V} + \mathcal{F}\mathcal{P}\mathcal{V})(\mathcal{T}\mathcal{P}\mathcal{V} + \mathcal{F}\mathcal{N}\mathcal{V})(\mathcal{T}\mathcal{N}\mathcal{V} + \mathcal{F}\mathcal{P}\mathcal{V})(\mathcal{T}\mathcal{N}\mathcal{V} + \mathcal{F}\mathcal{N}\mathcal{V})}} \quad (42)$$

$$\text{Kappa_Score} = \frac{P_o - P_e}{1 - p_e} \quad (43)$$

4.3. Results

Experiments are divided into five categories. The first category contains the results of the proposed methodology without optimizing the features, and the results will be taken on the fused extracted feature vector that is ϕ^{Fused} . In the second, third, and fourth experiments, the fused feature vector ϕ^{Fused} is further optimized using three different algorithms. The optimized feature vector obtained as the output is fed to the machine learning classifiers. We have evaluated these experiments on two datasets (the description of datasets is given in the previous section).

4.3.1. Experiment 1: Results on the ϕ^{Fused} Feature Vector

The ϕ^{Fused} feature vector is extracted by employing fusion on three different features that are extracted using: the ResNet_18 model, the proposed XcepNet23 model, and texture features extracted using the LBP method. We have performed this experiment on two datasets: the Hybrid_Dataset and the Kvasir_V1 dataset. After preprocessing the datasets, two types of CNN features and LBP features are extracted from the dataset. The obtained features of the Hybrid_Dataset returned a feature vector mathematically denoted as ϕ^{Fused} . The feature vectors obtained from the ResNet18 model and XcepNet23 are of $\mathcal{N} \times 512$ dimensions, where \mathcal{N} represents the Hybrid_Dataset's total number of images. A feature vector of $\mathcal{N} \times 59$ is obtained through the LBP. After the fusion of these three vectors, a fused feature vector is obtained, and the dimensions of $\phi^{Fused-1}$ are $\mathcal{N} \times 1083$; these features are used by classifiers to identify and classify GIT abnormalities. Q_SVM has achieved 100% accuracy in 124.94 s. In terms of the other performance metrics, such as the SEN, SPE, PRE, and F1 score, Q_SVM performed well. E_S_Disc achieved the second-highest accuracy of 99.98% in 483.66 s, and F_Tree achieved the third-highest accuracy of 99.86% in 32.840 s. C_Tree had the worst classification performance, with an accuracy of 82.39%. Table 1 displays the quantitative findings of this experiment.

Similarly, preprocessed Kvasir-V1 images are utilized for feature extraction. The feature vectors obtained from the ResNet18 model and XcepNet23 are of $\mathcal{M} \times 512$ dimensions, where \mathcal{M} symbolizes the total number of images in the Hybrid_Dataset. A feature vector of $\mathcal{M} \times 59$ is obtained through the LBP. In the next step, feature fusion is performed to obtain a strong set of features. The ensemble feature vector obtained from the Kvasir_V1 dataset is of dimensions $\mathcal{M} \times 1083$. A fused feature vector is produced when these three extracted feature vectors are an ensemble that is denoted by $\phi^{Fused-2}$; the ensemble feature vector is

utilized for the classification. The highest classification accuracy achieved on the Kvasir_V1 dataset is 98.60% on the Q_SVM classifier in 41.758 s. With 98.47%, E_S_Disc achieved the second-highest accuracy, while C_KNN came in third with a 97.97% accuracy. The results of this experiment are depicted graphically in Figure 8 for both datasets.

Table 1. Results on the ϕ^{Fused} feature vector.

Dataset Name	Classifier	ACC (%)	SEN (%)	PRE (%)	SPE (%)	F1 (%)	Kappa Score	MCC	Time (s)
Hybrid_Dataset	F_Tree	99.96	99.96	99.96	99.99	99.96	1.000	1.000	32.840
	C_Tree	82.39	75.45	87.64	95.66	70.22	0.75	0.790	21.435
	Q_SVM	100.0	100.0	100.0	100.0	100.0	1.000	1.000	124.94
	F_G_SVM	91.26	89.04	95.34	97.54	91.25	1.000	1.000	1286.8
	C_KNN	99.76	99.76	99.74	99.94	99.75	1.000	1.000	786.61
	E_S_Disc	99.98	99.98	99.98	99.99	99.98	1.000	1.000	483.66
Kvasir_V1	F_Tree	91.50	91.49	91.57	98.78	91.52	1.000	1.000	20.989
	C_Tree	59.45	59.45	65.68	94.2	55.35	0.7142	0.7412	17.428
	Q_SVM	98.60	98.60	98.60	99.80	98.60	1.000	1.000	41.758
	F_G_SVM	37.17	37.17	76.83	91.02	37.32	0.000	0.000	232.78
	C_KNN	97.97	97.97	98.06	99.71	97.99	1.000	1.000	56.647
	E_S_Disc	98.47	98.47	98.47	99.78	98.47	1.000	1.000	173.25

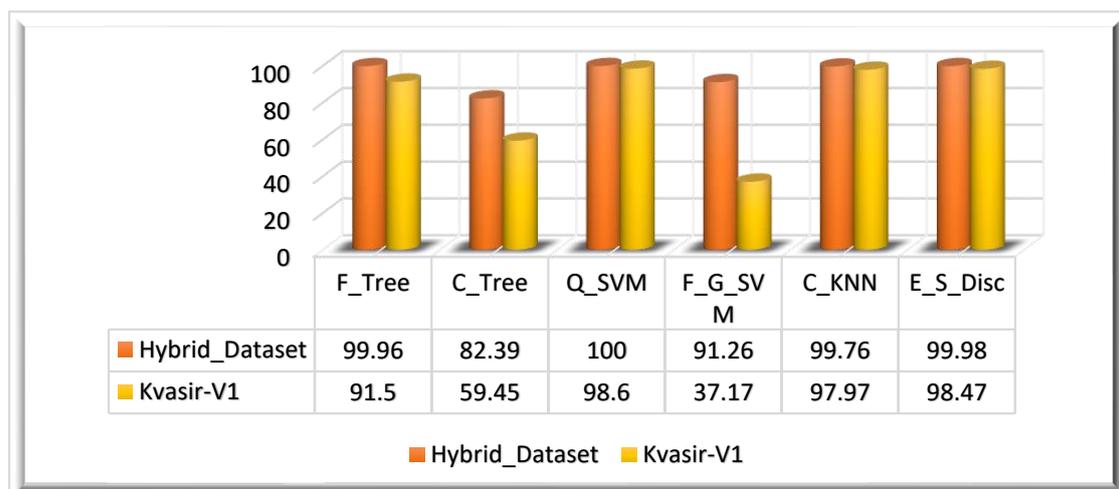


Figure 8. Experiment 1 results.

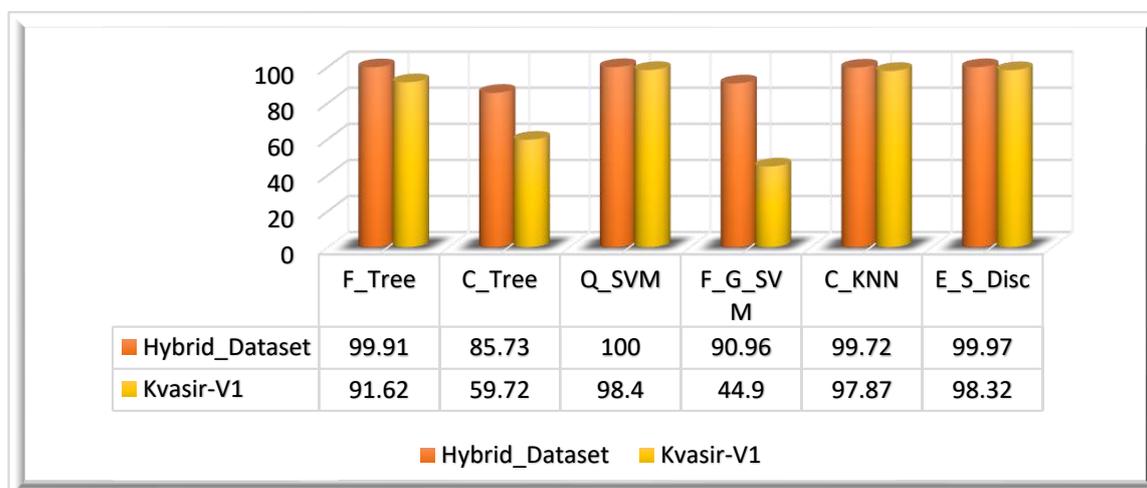
4.3.2. Experiment 2: Results on the ϕ_{BDA}^{Fused} Feature Vector

The ensemble feature vectors $\phi^{Fused-1}$ and $\phi^{Fused-2}$ are optimized using the BDA algorithm. In this experiment, the results are taken on the ϕ_{BDA}^{Fused} feature vector. For Hybrid_Dataset, the BDA algorithm returned a feature vector with the dimensions $\mathcal{N} \times 384$. The optimized feature vector is utilized for the classification of the dataset. For the evaluation of BDA-optimized features, we have utilized multiple classifiers. From the results, it can be observed that Q_SVM has achieved 100% accuracy in 58.140 s. Q_SVM performed well compared to the other classifiers. E_S_Disc achieved the second-highest accuracy of 99.97% in 114.66 s, and C_KNN achieved the third-highest accuracy of 97.72% in 264.70 s. C_Tree had the worst classification performance, with an accuracy of 85.73%. Table 2 shows the results of this experiment.

Table 2. Results on the ϕ_{BDA}^{Fused} feature vector.

Dataset Name	Classifier	ACC (%)	SEN (%)	PRE (%)	SPE (%)	F1 (%)	Kappa Score	MCC	Time (s)
Hybrid_Dataset	F_Tree	99.91	99.91	99.91	99.97	99.91	1.000	1.000	14.220
	C_Tree	85.73	78.75	90.95	96.45	73.94	0.750	0.791	7.8833
	Q_SVM	100.0	100.0	100.0	100.0	100.0	1.000	1.000	58.140
	F_G_SVM	90.96	88.67	95.22	97.46	90.93	1.000	1.000	506.73
	C_KNN	99.72	99.73	99.70	99.93	99.71	1.000	1.000	264.70
	E_S_Disc	99.97	99.97	99.97	99.99	99.97	1.000	1.000	114.66
Kvasir_V1	F_Tree	91.62	91.62	91.82	98.80	91.71	1.000	1.000	12.289
	C_Tree	59.72	59.72	77.70	94.24	47.16	0.571	0.604	11.308
	Q_SVM	98.40	98.40	98.37	99.77	98.38	1.000	1.000	22.685
	F_G_SVM	44.90	44.90	73.70	92.12	42.78	0.428	0.495	92.103
	C_KNN	97.87	97.89	97.97	99.69	97.92	1.000	1.000	28.063
	E_S_Disc	98.32	98.32	98.32	99.76	98.32	1.000	1.000	66.739

In this experiment, we have also computed the ϕ_{BDA}^{Fused} feature vector on the Kvasir_V1 dataset. The ensemble feature vector obtained from the Kvasir_V1 dataset is of $\mathcal{M} \times 1083$ dimensions. The ensemble feature vector is optimized using the BDA algorithm. The optimized feature vector has $\mathcal{M} \times 499$ dimensions. The highest classification accuracy of this experiment (Kvasir_V1 dataset) is 98.40%, achieved on the Q_SVM classifier in 22.685 s. On the Kvasir_V1 dataset, the E_S_Disc classifier achieved the second-highest accuracy of 98.32% in 66.739 s, and C_KNN achieved the thirteenth-highest accuracy of 97.87%. F_G_SVM performed worst among all classifiers, achieving a 44.90% accuracy in 92.103 s. The visual representation of the results is given in Figure 9 for both datasets.

**Figure 9.** Experiment 2 results.

4.3.3. Experiment 3: Results on the ϕ_{MFO}^{Fused} Feature Vector

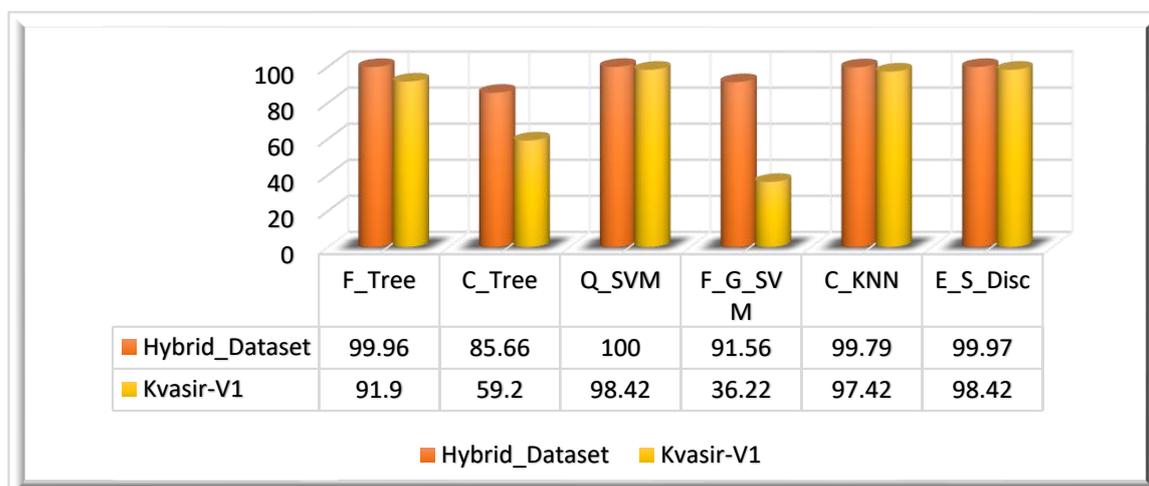
Experiment 3 is based on the ϕ_{MFO}^{Fused} feature vector. The ensemble feature vectors $\phi^{Fused-1}$ and $\phi^{Fused-2}$ are optimized using the MFO algorithm; this algorithm returned a feature vector with the dimensions $\mathcal{N} \times 437$ for the Hybrid_Dataset. Table 3 presents the quantitative results of this experiment. For the evaluation of MFO-optimized features, we have utilized multiple classifiers. From the results, it can be observed that Q_SVM has achieved 100% accuracy in 67.941 s. E_S_Disc, one of the classifiers that was assessed, came in second, with an accuracy rating of 99.97%, trailing only F_Tree (99.76%). The least effective model, C_Tree, has a classification accuracy of 85.66%.

Table 3. Results on the ϕ_{MFO}^{Fused} feature vector.

Dataset Name	Classifier	ACC (%)	SEN (%)	PRE (%)	SPE (%)	F1 (%)	Kappa Score	MCC	Time (s)
Hybrid_Dataset	F_Tree	99.96	99.96	99.96	99.99	99.96	1.000	1.000	11.559
	C_Tree	85.66	78.68	90.82	96.44	73.85	0.750	0.791	11.000
	Q_SVM	100.0	100.0	100.0	100.0	100.0	1.000	1.000	67.941
	F_G_SVM	91.56	89.31	95.46	97.63	91.45	1.000	1.000	563.31
	C_KNN	99.79	99.8	99.76	99.94	99.78	1.000	1.000	302.73
	E_S_Disc	99.97	99.97	99.97	99.99	99.97	1.000	1.000	135.34
Kvasir_V1	F_Tree	91.90	91.90	92.02	98.84	91.95	1.000	1.000	13.461
	C_Tree	59.20	59.20	74.32	94.17	50.62	0.571	0.625	5.6968
	Q_SVM	98.37	98.37	98.35	99.76	98.36	1.000	1.000	36.044
	F_G_SVM	36.22	36.22	77.41	90.88	36.23	0.000	0.000	102.24
	C_KNN	97.42	97.45	97.58	99.63	97.47	1.000	1.000	28.223
	E_S_Disc	98.42	98.42	98.45	99.77	98.43	1.000	1.000	60.429

Experiment 3 is also evaluated on the Kvasir_V1 dataset. The MFO-optimized feature vector ϕ_{MFO}^{Fused} is obtained by employing the MFO algorithm on the fused feature vector ($\phi^{Fused-2}$). The ensemble feature vector optimized using the MFO algorithm returned a feature vector of $\mathcal{M} \times 529$ dimensions. The optimized feature vector ϕ_{MFO}^{Fused} is classified using multiple classifiers. The highest classification accuracy of this experiment (Kvasir_V1 dataset) is 98.42%, achieved on the E_S_Disc classifier in 60.429 s.

On the Kvasir_V1 dataset, the Q_SVM classifier achieved the second-highest accuracy of 98.37% in 36.044 s, and C_KNN achieved the thirteenth highest accuracy of 97.42%. F_G_SVM performed worst among all classifiers by achieving a 36.22% accuracy in 102.24 s. The graphical representation of the results is given in Figure 10 for both datasets.

**Figure 10.** Experiment 3 results.

4.3.4. Experiment 4: Results on the ϕ_{PSO}^{Fused} Feature Vector

In Experiment 4, we utilized the PSO algorithm for feature optimization; it is primarily reliant on the ϕ_{PSO}^{Fused} feature vector. The ensemble feature vector ϕ^{Fused} is optimized using the PSO algorithm; this algorithm returned a feature vector with the dimensions $\mathcal{N} \times 428$. The optimal set of features is utilized in this experiment for the evaluation of the PSO optimization algorithm performance. For the evaluation of PSO-optimized features, we have utilized multiple classifiers. Similar to other experiments, the performance is evaluated using the same performance measures such as the ACC, SEN, PRE, SPE, F1-Score, and execution time. The results depict that Q_SVM has achieved 100% accuracy in

64.238 s. E_S_Disc came in second, with an accuracy of 99.93%, and F_Tree achieved the third-highest accuracy of 99.92%. With an accuracy of 85.81%, C_Tree performed the worst in terms of classification. Table 4 displays the experiment's findings.

Table 4. Results on the ϕ_{PSO}^{Fused} feature vector.

Dataset Name	Classifier	ACC (%)	SEN (%)	PRE (%)	SPE (%)	F1 (%)	Kappa Score	MCC	Time (s)
Hybrid_Dataset	F_Tree	99.92	99.92	99.91	99.98	99.92	1.000	1.000	13.431
	C_Tree	85.81	78.87	91.13	96.47	73.98	0.750	0.791	9.9464
	Q_SVM	100.0	100.0	100.0	100.0	100.0	1.000	1.000	64.238
	F_G_SVM	91.55	89.26	95.46	97.62	91.4	1.000	1.000	560.76
	C_KNN	99.77	99.77	99.75	99.94	99.76	1.000	1.000	296.58
	E_S_Disc	99.93	99.93	99.93	99.98	99.93	1.000	1.000	132.16
Kvasir_V1	F_Tree	91.15	91.15	91.21	98.73	91.17	1.000	1.000	11.495
	C_Tree	58.82	58.82	64.93	94.11	53.31	0.704	0.7313	7.3797
	Q_SVM	99.24	98.69	98.69	99.87	98.69	1.000	1.000	23.957
	F_G_SVM	39.07	39.07	78.32	91.29	39	0.142	0.285	92.005
	C_KNN	97.62	97.62	97.71	99.66	97.65	1.000	1.000	25.571
	E_S_Disc	98.45	98.45	98.47	99.77	98.46	1.000	1.000	66.56

Experiment 3 is also evaluated on the Kvasir_V1 dataset. The PSO-optimized feature vector ϕ_{PSO}^{Fused} is obtained by employing the PSO algorithm on the ensemble feature vector. The ensemble feature vector optimized using the PSO algorithm returned a feature vector of $\mathcal{M} \times 525$ dimensions. The highest classification accuracy of this experiment (Kvasir_V1 dataset) is 99.24%, achieved on Q_SVM classifiers in 23.957 s. On the Kvasir_V1 dataset, the E_S_Disc classifier achieved the second-highest accuracy of 98.45% in 66.56 s, and C_KNN achieved the thirteenth-highest accuracy of 97.62%. F_G_SVM performed worst among all classifiers by achieving a 39.07% accuracy in 229.41 s. The visual representation of the results is given in Figure 11 for both datasets.

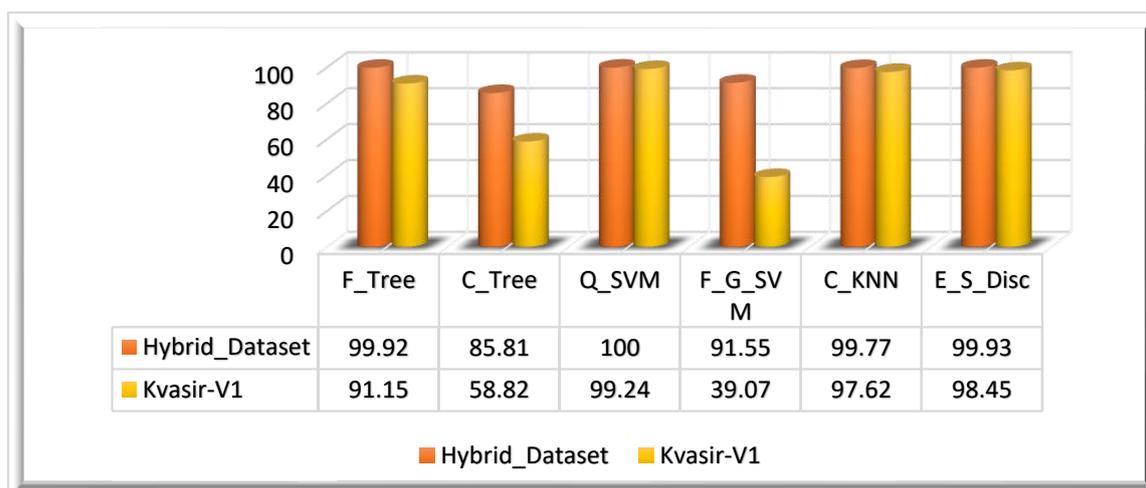


Figure 11. Experiment 4 results.

5. Discussion and Comparison with Existing Methods

The proposed approach is compared with the most recent methods in this section [1,3,19,33,35,42,43,56,57]. The authors in [3] proposed a method for gastrointestinal disease detection from WCE images, and they achieved a 98.40% accuracy. In [3], Khan et al. achieved a 98.40% accuracy. The methodology proposed in [35] is evaluated on the three datasets. This method achieved 99.25%, 99.90%, and 100.0% accuracies on the Kvasir-V1, Nerthus, and CUI WAH WCE datasets. In another research study [19], two datasets were

utilized for the assessment of the proposed method; the considered datasets were Kvasir-V1 and CUIWAH WCE. The achieved accuracy on the CUI WAH WCE dataset was 99.80%, whereas on Kvasir-V1, an 87.80% accuracy was obtained. In [33], the authors achieved a 99.80% accuracy for the WCE image classification. Khan et al. [43] utilized the Hybrid dataset comprising 15,000 images and achieved a 99.50% accuracy. In [56,57], the authors proposed a method for classifying gastrointestinal disease detection and classification; they achieved 97% and 96.46% accuracies on the Kvasir-V1 dataset, respectively. Later on, the authors in [42] proposed a novel method for the GIT disease classification using two optimization algorithms, the proposed technique is named Moth-Crow-based features optimization. This approach outperformed the most recent methods on three datasets. This study used three datasets—CUI WAH WCE, Kvasir-V1, and Kvasir-V2—and achieved 99.42%, 97.85%, and 97.20% accuracies, respectively. In this research, we have utilized two datasets for the evaluation of the proposed method, which are the Hybrid Dataset (CUI WAH WCE, Kvasir-V1, CVC-Clinic) and the Kvasir-V1 dataset. Our contribution to this research is in the preprocessing phase, which returned a better set of features, which results in better performance. We have achieved 100% accuracy on the Hybrid Dataset and 99.24% accuracy on the Kvasir-V1 dataset in 58.140 and 23.957 s respectively.

Table 5 provides a comparison of the results obtained by the proposed method and existing methods. It is observed from the comparison that our proposed method has performed good on the Kvasir-V1 dataset in terms of all the performance measures, if we compare the performance with the results of state-of-the-art research studies [19,42,56,57]. One of our previous studies achieved a 99.25% accuracy, which is 0.01% better than that of the present research due to the complexity of this method. The proposed method achieved the results in less time as compared to the previous method. This is the limitation of the proposed method that can be addressed in a future research study by using the more advanced algorithms. In this research, we have utilized a hybrid dataset that comprises the three different dataset images. The proposed method has achieved 100% accuracy on the hybrid dataset in 58.140 s, which is a good achievement. It is evident from this comparison that our proposed approach has outperformed the most recent ones in terms of performance.

Table 5. Comparison with Existing Techniques.

Ref. No.	Year	Dataset	Total Images	Dataset Classes	ACC (%)	SEN (%)	PRE (%)	SPE (%)	F1-Score	Time (s)
[3]	2020	CUI WAH WCE	6000	3	98.40	98.33	98.36	-	98.34	-
[1]	2021	CUI WAH WCE	4500	3	99.54	100.0	99.51	-	-	31.700
[19]	2021	Kvasir-V1	4000	8	87.80	87.40	87.99	98.06	87.63	50.700
		CUI WAH WCE	2326	3	99.80	99.83	99.80	99.92	99.81	17.031
[33]	2021	CUI WAH WCE	4000	4	99.80	99.00	99.25	-	99.12	211.90
[35]	2021	Kvasir-V1	4000	8	99.25	99.30	99.89	99.25	-	24.958
		Nerthus	5524	4	99.90	99.75	99.84	99.94	-	35.033
		CUI WAH WCE	2326	3	100.0	100.0	100.0	100.0	-	7.5840
[43]	2021	Hybrid Dataset	15,000	5	99.50	99.50	96.00	-	97.70	48.830

Table 5. Cont.

Ref. No.	Year	Dataset	Total Images	Dataset Classes	ACC (%)	SEN (%)	PRE (%)	SPE (%)	F1-Score	Time (s)
[56]	2021	Kvasir-V1	4000	8	97.00	96.80	-	99.20	-	-
[57]	2022	Kvasir-V1	4000	8	96.46	94.46	94.46	-	94.46	-
[42]	2022	CUI WAH WCE	4000	8	99.42	-	-	-	-	91.013
		Kvasir-V1	4000	8	97.85	-	-	-	-	71.031
		Kvasir-V2	8000	8	97.20	-	-	-	-	111.13
[58]	2022	Kvasir-V1	4000	8	98.20	-	-	-	-	52.504
		Kvasir-V2	8000	8	98.02	-	-	-	-	102.502
		CUI WAH WCE	-	3	99.61	-	-	-	-	69.544
Proposed		Hybrid Dataset	15,448	5	100.0	100.0	100.0	100.0	100.0	58.140
		Kvasir-V1	4000	8	99.24	98.69	98.69	99.87	98.69	23.957

6. Conclusions

This research proposes a deep learning model of 23 layers named XcepNet23 for Gastrointestinal disease detection and classification. We have proposed a hybrid image preprocessing framework that is employed on the augmented dataset. The proposed hybrid contrast stretching-based image enhancement is employed on the dataset. We have extracted the two types of CNN features in this research work. The first CNN feature vector is obtained using the Global Average pool layer of the proposed XcepNet23 model. The second CNN feature vector is obtained from the original augmented dataset by utilizing the fine-tuned ResNet18 model. The texture variation of the region of interest is a big challenge in the accurate recognition of gastrointestinal diseases. In this research, we have extracted the texture features using the LBP method. The obtained feature vectors are ensemble to obtain the strong set of features that incorporate the two types of CNN and texture features. The fused feature vector is optimized using three different nature-inspired meta heuristic optimization algorithms (BDA, MFO, PSO). The experiments are conducted on two datasets: the Hybrid dataset and the Kvasir-V1 dataset. The evaluation of the method is performed on three different feature vectors: the original ensemble feature vector, BDA-optimized ensemble feature vector, MFO-optimized ensemble feature vector, and PSO-optimized ensemble feature vector. In light of the results of the comparison between the proposed method and the techniques that were already proposed, it is concluded that our method has performed better across all performance metrics. This comparison demonstrates that our proposed strategy has performed better than the most recent ones. In 58.140 and 23.957 s, respectively, we were able to achieve 100% accuracy on the Hybrid Dataset and 99.24% accuracy on the Kvasir-V1 dataset.

The only limitation of this research work is the segmentation of the gastrointestinal tract abnormalities. The most noticed abnormalities are bleeding, ulcers, and polyps. The segmentation of bleeding, ulcers, and polyps can be performed along with the recognition of diseases in future research work. It is difficult to detect three different types of diseases using the same methodology, as these three types of abnormalities have different color, texture, and shape variations. The algorithm can be expanded in future research to identify and categorize a wider variety of gastrointestinal disorders. This might entail gathering more labeled data for less prevalent diseases or researching transfer learning strategies to use information from similar disorders. To implement the approach in clinics and enable real-time prediction, methods can also be researched. Moreover, in future work, this method may be improved by reducing the computational complexity.

Author Contributions: Conceptualization and Methodology, J.N.; Analysis M.I.S. (Muhammad Imran Sharif) and M.I.S. (Muhammad Irfan Sharif); Validation, H.T.R. and S.K.; Formal analysis, A.E.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by King Saud University for funding this work through Researchers Supporting Project number (RSPD2023R711), King Saud University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Publically available datasets are utilized in this research.

Data Availability Statement: Dataset utilized in this research is publically available.

Acknowledgments: The authors extend their appreciation to King Saud University for funding this work through Researchers Supporting Project number (RSPD2023R711), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Sharif, M.; Khan, M.A.; Rashid, M.; Yasmin, M.; Afza, F.; Tanik, U.J. Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. *J. Exp. Theor. Artif. Intell.* **2019**, *33*, 577–599. [[CrossRef](#)]
- Mohammad, F.; Al-Razgan, M. Deep Feature Fusion and Optimization-Based Approach for Stomach Disease Classification. *Sensors* **2022**, *22*, 2801. [[CrossRef](#)]
- Khan, M.A.; Kadry, S.; Alhaisoni, M.; Nam, Y.; Zhang, Y.-D.; Rajinikanth, V.; Sarfaraz, M.S. Computer-Aided Gastrointestinal Diseases Analysis from Wireless Capsule Endoscopy: A Framework of Best Features Selection. *IEEE Access* **2020**, *8*, 132850–132859. [[CrossRef](#)]
- Kourie, H.R.; Tabchi, S.; Ghosn, M. Checkpoint inhibitors in gastrointestinal cancers: Expectations and reality. *World J. Gastroenterol.* **2017**, *23*, 3017–3021. [[CrossRef](#)] [[PubMed](#)]
- Long, J.; Lin, J.; Wang, A.; Wu, L.; Zheng, Y.; Yang, X.; Wan, X.; Xu, H.; Chen, S.; Zhao, H. PD-1/PD-L blockade in gastrointestinal cancers: Lessons learned and the road toward precision immunotherapy. *J. Hematol. Oncol.* **2017**, *10*, 146. [[CrossRef](#)] [[PubMed](#)]
- Umar, S.B.; Fleischer, D.E. Esophageal cancer: Epidemiology, pathogenesis and prevention. *Nat. Clin. Pract. Gastroenterol. Hepatol.* **2008**, *5*, 517–526. [[CrossRef](#)]
- Pennathur, A.; Gibson, M.K.; Jobe, B.A.; Luketich, J.D. Oesophageal carcinoma. *Lancet* **2013**, *381*, 400–412. [[CrossRef](#)] [[PubMed](#)]
- Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
- Sivakumar, P.; Kumar, B.M. A novel method to detect bleeding frame and region in wireless capsule endoscopy video. *Clust. Comput.* **2019**, *22*, 12219–12225. [[CrossRef](#)]
- Amin, J.; Sharif, M.; Yasmin, M.; Fernandes, S.L. A distinctive approach in brain tumor detection and classification using MRI. *Pattern Recognit. Lett.* **2020**, *139*, 118–127. [[CrossRef](#)]
- Amin, J.; Sharif, M.; Yasmin, M.; Ali, H.; Fernandes, S.L. A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions. *J. Comput. Sci.* **2017**, *19*, 153–164. [[CrossRef](#)]
- Liaqat, A.; Khan, M.A.; Shah, J.H.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Automated ulcer and bleeding classification from wce images using multiple features fusion and selection. *J. Mech. Med. Biol.* **2018**, *18*, 1850038. [[CrossRef](#)]
- Lee, N.M.; Eisen, G.M. 10 years of capsule endoscopy: An update. *Expert Rev. Gastroenterol. Hepatol.* **2010**, *4*, 503–512. [[CrossRef](#)] [[PubMed](#)]
- Saeed, T.; Loo, C.K.; Kassim, M.S.S. Ensembles of Deep Learning Framework for Stomach Abnormalities Classification. *Comput. Mater. Contin.* **2022**, *70*, 4357–4372. [[CrossRef](#)]
- Rajinikanth, V.; Satapathy, S.C.; Fernandes, S.L.; Nachiappan, S. Entropy based segmentation of tumor from brain MR images—A study with teaching learning based optimization. *Pattern Recognit. Lett.* **2017**, *94*, 87–95. [[CrossRef](#)]
- Rajinikanth, V.; Fernandes, S.L.; Bhushan, B.; Harisha, S.; Sunder, N.R. Segmentation and analysis of brain tumor using Tsallis entropy and regularised level set. In *Proceedings of the 2nd International Conference on Micro-Electronics, Electromagnetics and Telecommunications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 313–321.
- Bokhari, S.T.F.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Fundus Image Segmentation and Feature Extraction for the Detection of Glaucoma: A New Approach. *Curr. Med. Imaging Former. Curr. Med. Imaging Rev.* **2018**, *14*, 77–87. [[CrossRef](#)]
- Naqi, S.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Lung Nodule Detection Using Polygon Approximation and Hybrid Features from CT Images. *Curr. Med. Imaging Former. Curr. Med. Imaging Rev.* **2018**, *14*, 108–117. [[CrossRef](#)]
- Naz, J.; Khan, M.A.; Alhaisoni, M.; Song, O.-Y.; Tariq, U.; Kadry, S. Segmentation and Classification of Stomach Abnormalities Using Deep Learning. *Comput. Mater. Contin.* **2021**, *69*, 607–625. [[CrossRef](#)]
- Nida, N.; Sharif, M.; Khan, M.U.G.; Yasmin, M.; Fernandes, S.L. A framework for automatic colorization of medical imaging. *IIOAB J.* **2016**, *7*, 202–209.

21. Suman, S.; Hussin, F.A.; Malik, A.S.; Walter, N.; Goh, K.L.; Hilmi, I. Image enhancement using geometric mean filter and gamma correction for WCE images. In Proceedings of the International Conference on Neural Information Processing, Kuching, Malaysia, 3–6 November 2014; Springer: Berlin/Heidelberg, Germany; pp. 276–283.
22. Rajinikanth, V.; Madhavaraja, N.; Satapathy, S.C.; Fernandes, S.L. Otsu's multi-thresholding and active contour snake model to segment dermoscopy images. *J. Med. Imaging Health Inform.* **2017**, *7*, 1837–1840. [[CrossRef](#)]
23. Lee, Y.-G.; Yoon, G. Real-time image analysis of capsule endoscopy for bleeding discrimination in embedded system platform. *Int. J. Biomed. Biol. Eng.* **2011**, *5*, 583–587.
24. Yuan, Y.; Li, B.; Meng, M.Q.-H. WCE Abnormality Detection Based on Saliency and Adaptive Locality-Constrained Linear Coding. *IEEE Trans. Autom. Sci. Eng.* **2016**, *14*, 149–159. [[CrossRef](#)]
25. Fu, Y.; Zhang, W.; Mandal, M.; Meng, M.Q.-H. Computer-aided bleeding detection in WCE video. *IEEE J. Biomed. Health Inform.* **2013**, *18*, 636–642. [[CrossRef](#)]
26. Mathew, M.; Gopi, V.P. Transform based bleeding detection technique for endoscopic images. In Proceedings of the 2015 2nd International Conference on Electronics and Communication Systems (ICECS), Coimbatore, India, 26–27 February 2015; IEEE: Piscataway, NJ, USA; pp. 1730–1734.
27. Yuan, Y.; Meng, M.Q.-H. Polyp classification based on bag of features and saliency in wireless capsule endoscopy. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; IEEE: Piscataway, NJ, USA; pp. 3930–3935.
28. Ghosh, T.; Bashar, S.K.; Alam, S.; Wahid, K.; Fattah, S.A. A statistical feature based novel method to detect bleeding in wireless capsule endoscopy images. In Proceedings of the 2014 International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, Bangladesh, 23–24 May 2014; IEEE: Piscataway, NJ, USA; pp. 1–4.
29. Yuan, Y.; Meng, M.Q.-H. Deep learning for polyp recognition in wireless capsule endoscopy images. *Med. Phys.* **2017**, *44*, 1379–1389. [[CrossRef](#)]
30. Yuan, Y.; Wang, J.; Li, B.; Meng, M.Q.-H. Saliency Based Ulcer Detection for Wireless Capsule Endoscopy Diagnosis. *IEEE Trans. Med. Imaging* **2015**, *34*, 2046–2057. [[CrossRef](#)]
31. Suman, S.; Hussin, F.A.; Malik, A.S.; Ho, S.H.; Hilmi, I.; Leow, A.H.-R.; Goh, K.-L. Feature Selection and Classification of Ulcerated Lesions Using Statistical Analysis for WCE Images. *Appl. Sci.* **2017**, *7*, 1097. [[CrossRef](#)]
32. Sutton, R.T.; Zaiane, O.R.; Goebel, R.; Baumgart, D.C. Artificial intelligence enabled automated diagnosis and grading of ulcerative colitis endoscopy images. *Sci. Rep.* **2022**, *12*, 2748. [[CrossRef](#)] [[PubMed](#)]
33. Nayyar, Z.; Khan, M.A.; Alhusein, M.; Nazir, M.; Aurangzeb, K.; Nam, Y.; Kadry, S.; Haider, S.I. Gastric Tract Disease Recognition Using Optimized Deep Learning Features. *Comput. Mater. Contin.* **2021**, *68*, 2041–2056. [[CrossRef](#)]
34. Zhang, G.; Pan, J.; Xing, C. Computer-aided diagnosis of digestive tract tumor based on deep learning for medical images. *Netw. Model. Anal. Health Inform. Bioinform.* **2022**, *11*, 8. [[CrossRef](#)]
35. Naz, J.; Sharif, M.; Raza, M.; Shah, J.H.; Yasmin, M.; Kadry, S.; Vimal, S. Recognizing Gastrointestinal Malignancies on WCE and CCE Images by an Ensemble of Deep and Handcrafted Features with Entropy and PCA Based Features Optimization. *Neural Process. Lett.* **2021**, *55*, 115–140. [[CrossRef](#)]
36. Liu, Y. Noise reduction by vector median filtering. *Geophysics* **2013**, *78*, V79–V87. [[CrossRef](#)]
37. Chien, C.-L.; Tseng, D.-C. Color image enhancement with exact HSI color model. *Int. J. Innov. Comput. Inf. Control.* **2011**, *7*, 6691–6710.
38. Ibraheem, N.A.; Hasan, M.M.; Khan, R.Z.; Mishra, P.K. Understanding color models: A review. *ARNP J. Sci. Technol.* **2012**, *2*, 265–275.
39. Saalfeld, S. CLAHE (Contrast Limited Adaptive Histogram Equalization). 2009. Available online: <https://imagej.nih.gov/ij/plugins/clahe/index.html> (accessed on 8 June 2023).
40. Ojala, T.; Pietikainen, M.; Harwood, D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, Israel, 9–13 October 1994; IEEE Computer Society Press: Piscataway, NJ, USA, 1994; pp. 582–585.
41. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
42. Khan, M.A.; Muhammad, K.; Wang, S.H.; Alsubai, S.; Binbusayyis, A.; Alqahtani, A.; Majumdar, A.; Thinnukool, O. Gastrointestinal diseases recognition: A framework of deep neural network and improved moth-crow optimization with DCCA fusion. *Hum.-Cent. Comput. Inf. Sci.* **2022**, *12*, 25.
43. Khan, M.A.; Majid, A.; Hussain, N.; Alhaisoni, M.; Zhang, Y.-D.; Kadry, S.; Nam, Y. Multiclass Stomach Diseases Classification Using Deep Learning Features Optimization. *Comput. Mater. Contin.* **2021**, *67*, 3381–3399. [[CrossRef](#)]
44. Khan, M.A.; Khan, M.A.; Ahmed, F.; Mittal, M.; Goyal, L.M.; Hemanth, D.J.; Satapathy, S.C. Gastrointestinal diseases segmentation and classification based on duo-deep architectures. *Pattern Recognit. Lett.* **2020**, *131*, 193–204. [[CrossRef](#)]
45. Too, J.; Mirjalili, S. A Hyper Learning Binary Dragonfly Algorithm for Feature Selection: A COVID-19 Case Study. *Knowl.-Based Syst.* **2021**, *212*, 106553. [[CrossRef](#)]
46. Mirjalili, S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Comput. Appl.* **2016**, *27*, 1053–1073. [[CrossRef](#)]

47. Mirjalili, S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowl. Based Syst.* **2015**, *89*, 228–249. [[CrossRef](#)]
48. Eberhart, R.; Kennedy, J. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
49. Brieman, L.; Friedman, J.; Stone, C.J.; Olshen, R. *Classification and Regression Tree Analysis*; CRC Press: Boca Raton, FL, USA, 1984.
50. Priyam, A.; Abhijeeta, G.R.; Rathee, A.; Srivastava, S. Comparative analysis of decision tree classification algorithms. *Int. J. Curr. Eng. Technol.* **2013**, *3*, 334–337.
51. Zhong, W.; He, G.; Pi, D.; Sun, Y. SVM with quadratic polynomial kernel function based nonlinear model one-step-ahead predictive control. *Chin. J. Chem. Eng.* **2005**, *13*, 373–379.
52. Li, Y.; Tian, X.; Song, M.; Tao, D. Multi-task proximal support vector machine. *Pattern Recognit.* **2015**, *48*, 3249–3257. [[CrossRef](#)]
53. Xu, Y.; Zhu, Q.; Fan, Z.; Qiu, M.; Chen, Y.; Liu, H. Coarse to fine K nearest neighbor classifier. *Pattern Recognit. Lett.* **2013**, *34*, 980–986. [[CrossRef](#)]
54. Radhika, K.; Varadarajan, S. *Ensemble Subspace Discriminant Classification of Satellite Images*; NISCAIR-CSIR: Delhi, India, 2018.
55. Pogorelov, K.; Randel, K.R.; Griwodz, C.; Eskeland, S.L.; de Lange, T.; Johansen, D.; Spampinato, C.; Dang-Nguyen, D.-T.; Lux, M.; Schmidt, P.T.; et al. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In Proceedings of the 8th ACM on Multimedia Systems Conference, New York, NY, USA, 20–23 June 2017; pp. 164–169.
56. Al-Adhaileh, M.H.; Senan, E.M.; Alsaade, F.W.; Aldhyani, T.H.H.; Alsharif, N.; Alqarni, A.A.; Uddin, M.I.; Alzahrani, M.Y.; Alzain, E.D.; Jadhav, M.E. Deep Learning Algorithms for Detection and Classification of Gastrointestinal Diseases. *Complexity* **2021**, *2021*, 6170416. [[CrossRef](#)]
57. Kumar, C.A.; Mubarak, D. Classification of Early Stages of Esophageal Cancer Using Transfer Learning. *IRBM* **2022**, *43*, 251–258. [[CrossRef](#)]
58. Khan, M.A.; Sahar, N.; Khan, W.Z.; Alhaisoni, M.; Tariq, U.; Zayyan, M.H.; Kim, Y.J.; Chang, B. GastroNet: A Framework of Saliency Estimation and Optimal Deep Learning Features Based Gastrointestinal Diseases Detection and Classification. *Diagnostics* **2022**, *12*, 2718. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.